# course 6 proj part 2

*Viraj Bhalala*

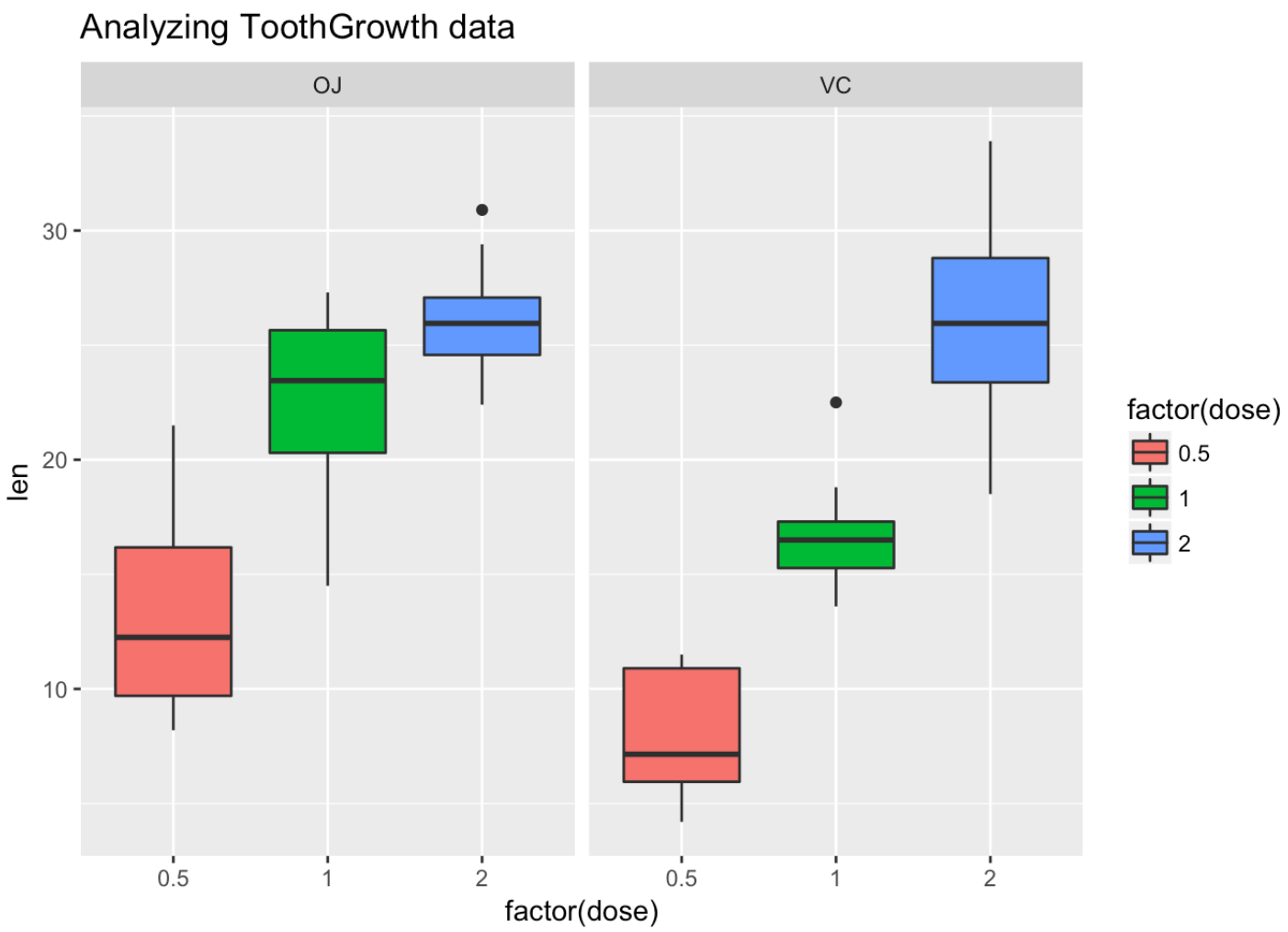
*April 23, 2017*

# Course 6 project part 2

we're going to analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Provide a basic summary of the data.
3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)
4. State your conclusions and the assumptions needed for your conclusions.

```
library(datasets)
data(ToothGrowth)
library(ggplot2)
ggplot(ToothGrowth, aes(x=factor(dose), y=len, fill=factor(dose)))+geom_boxplot()+fac
et_grid(.~supp)+ggtitle("Analyzing ToothGrowth data")
```




```
summary(ToothGrowth)
```

```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

```
xBar<-mean(ToothGrowth$len[1:30])
yBar<-mean(ToothGrowth$len[31:60])
xVar<-(sd(ToothGrowth$len[1:30]))^2
yVar<-(sd(ToothGrowth$len[31:60]))^2
q<-(((xVar+yVar)/30)^2)/((((xVar/30)^2)+((yVar/30)^2))/29)
t<-qt(0.975, q)
yBar -xBar + c(-1,1)*t*sqrt(xVar/30 + yVar/30)
```

```
## [1] -0.1710156  7.5710156
```

Perform Hypothesis test

```
t.test(len~supp, data=ToothGrowth, paired=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

Hypothesis test for subset a

```
a<-subset(ToothGrowth, dose==0.5)
b<-subset(ToothGrowth, dose==1.0)
c<-subset(ToothGrowth, dose==2.0)
t.test(len~supp, data=a, paired=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

Hypothesis test for subset b

```
t.test(len~supp, data=b, paired=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

Hypothesis test for subset c

```
t.test(len~supp, data=c, paired=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

# conclusion

By running this test, we founded that as tooth size increases, the doses tend to be higher